Documentation for the Patient Entitlement and Diagnosis Summary File (PEDSF)
July 17, 2013

FILENAME 1:   pedsf.cancer.txt

LRECL 1:   4,937

FILENAME 2:   pedsf.5percent.txt (if requested)

LRECL 2:   2,426

SEER Cases diagnosed 1973-2009

PEDSF File created on June 30, 2013

SEER data extracted from November, 2012 SEER Submission

ABSTRACT: The PEDSF file contains one record per person for individuals in the SEER Program data base who have been matched with Medicare enrollment records. Of persons who were reported by the SEER Registries to have been diagnosed with cancer at age 65 or older, 94 percent were matched with Medicare enrollment records. For persons appearing in the PEDSF file, basic SEER Program diagnostic information is available for up to 10 diagnosed cancer occurrences. Data also include Medicare entitlement and utilization data from 1991-2011. Geographically based socioeconomic information from the U.S. Census Bureau has also been included using Census 1990 and Census 2000.

PROGRAMMER: Angela Meekins, Information Management Services, Inc.

CONTACT: SEER-Medicare Technical Support, Information Management Services, Inc.
Email: SEER-Medicare@imsweb.com

NOTE1: PEDSF includes SEER cases from Los Angeles and San Jose from 1988 thru 2009, although these registries did not become part of the SEER program until 1992. Cases diagnosed prior to 1992 do not contain stage data. The LA and SJ registries have allowed the release of the data prior to 1992 but have indicated that, Cancer incidence data prior to 1992 have been collected under Subcontract 050C-8709 through 050I-8709 with the Public Health Institute. The subcontract is supported by the California Department of Health Services as part of its statewide cancer reporting program, mandated by Health and Safety Code Section 103875 and 103885. The ideas and opinions expressed from using these data are those of the author, and no endorsement of the State of California, Department of Health Services or the Public Health Foundation is intended or should be inferred.

NOTE2: For all non-custom requests, patients have been selected if the cancer(s) requested appeared in the site array (siterkm1-10) and the associated year of diagnosis (yrdx1-10) fell in the year of diagnosis window requested.

Documentation for the Patient Entitlement and Diagnosis Summary File (PEDSF)
July 17, 2013

COL	FIELD	LENGTH	SOURCE	NOTES
1	Patient ID **(patient_id)**	10	Master File	
1	SEER Cases - Registry	2	SEER	First two digits in the patient ID. 02 = Connecticut (1973+) 20 = Detroit (1973+) 21 = Hawaii (1973+) 22 = Iowa (1973+) 23 = New Mexico (1973+) 25 = Seattle (1974+) 26 = Utah (1973+) 42 = Kentucky (2000+) 43 = Louisiana (2000+) 44 = New Jersey (2000+) 87 = Georgia 88 = California NOTE: Specific California and Georgia registries will be differentiated in the repeated registry codes (starting in column 2489, variables reg1-reg10)
3	Case Number	8	Assigned by IMS	Encrypted SEER Case Number
11	Five Pct Indicator **(fivepct)**	1	Master File	Indicates inclusion in the 5% Medicare sample. Y = Included N = Not included
12	Match Type **(mat_type)**	1	Match File	CMSs matching algorithm is available at IMS. A = Alphabetic Match B = SSN Match
13	Match Flag 1 **(yobflg1)**	1	Match File	Year of Birth flag for SSN and Alpha Match. Y = Agreement between SEER and Medicare N = Did not agree 0 (Zero) = Cases matched on SSN U = Unknown, Missing data
14	Match Flag 2 **(dobflg2)**	1	Match File	Day of Birth Match for SSN and Alpha Match (See Match Flag 1 for code definition)
15	Match Flag 3 **(yodflg3)**	1	Match File	Month and Year of Death for Alpha Match. (See Match Flag 1 for code definition)
16	Match Flag 4 **(sexflg4)**	1	Match File	Sex Flag for SSN and Alpha Match. (See Match Flag 1 for code definition)
17	Match Flag 5 **(lstflg5)**	1	Match File	Last Name Flag for SSN Match. (See Match Flag 1 for code definition)

COL	FIELD	LENGTH	SOURCE	NOTES
18	Match Flag 6 (fstflg6)	1	Match File	First Name Flag for SSN Match. (See Match Flag 1 for code definition)
19	Match Flag 7 (mobflg7)	1	Match File	Month of Birth Flag for SSN Match. (See Match Flag 1 for code definition)
20	Match Flag 8 (midflg8)	1	Match File	Middle Initial Flag for Alpha Match. (See Match Flag 1 for code definition)
21	Number of digits in SSN that matched (numdigit)	1	Match File	Alpha search only. 0-8 9 = SSN match U = Unknown; data missing A = digits transposed and matched
22	Date of Death Flag (dod_flg)	1	Created at IMS	Shows the level of agreement between SEER and MEDICARE on the patient's month of death. Dates of death after 12/2010 were treated as not dead for comparison purposes. 0 = Not dead by 12/2010. 1 = Dead, both files agree. 2 = Dead, off by 1-3 months. 3 = Dead, off by 4-6 months. 4 = Dead in MEDICARE only, 5 = Dead in SEER only. 6 = Dead but number of months could not be calculated due to the fact that month was missing for either SEER or MEDICARE.
23	Date of Birth Flag (dob_flg)	1	Created at IMS	Shows the level of agreement between SEER and MEDICARE on the patient's month of birth. 0 = Both files agree on birth date. 1 = Birth date off by 1-3 months. 2 = Birth date off by 4-6 months. 3 = Birth date off by 7-11 months. 4 = Birth date off by one year 5 = Birth date off by 13-23 months 6 = Birth date off by 2 years 7 = Birth date off by 25-35 months 8 = Birth date off by 3 years 9 = Birth date off by 37+ months Blank = Birth date is missing
24	Linkage Flag (linkflag)	1	Created at IMS	Indicates the linkage the patient is first associated with. 1 = 2005 linkage, earliest claim year is 1991 2 = 2008 linkage, earliest claim year is 1998 3 = 2010 linkage, earliest claim year is 2000 4 = 2012 linkage, earliest claim year is 2002

COL	FIELD	LENGTH	SOURCE	NOTES
25	Medicare Date of Death **(med_dodm, med_dodd, med_dody)**	8	EDB	MMDDYYYY or blanks for not dead. Reported to CMS by SSA – not date of last contact. Date complete through 12/31/11
33	Medicare Month of Birth **(birthm)**	2	EDB	MM, zeroes for invalid
37	Medicare Year of Birth **(birthyr)**	4	EDB	YYYY, zeroes for invalid
41	Sex **(m_sex)**	1	EDB	1 = Male 2 = Female
42	Race **(race)**	1	EDB	1 = White 2 = Black 3 = Other 4 = Asian 5 = Hispanic 6 = N. Am. Native 0 = Unknown
43	Original Reason for Entitlement **(rsncd1)**	1	EDB	0 = OASI (age) 1 = DIB (disability) 2 = ESRD 3 = DIB/ESRD This is the first entitlement reason located on EDB
44	Current Reason for Entitlement **(cur_ent)**	1	EDB	Blank = Not Enrolled 0 = OASI (age) 1 = DIB 2 = ESRD 3 = DIB/ESRD Variable is selected by the first diagnosis date at age 65 or older, or the last diagnosis if all are prior to turning age 65. If information not available at those times then the first available data is used.
45	Current Reason for Entitlement Year **(cur_yr)**	4	EDB	YYYY

COL	FIELD	LENGTH	SOURCE	NOTES
49	Chronic Renal Disease **(chr_esrd)**	1	EDB	0 (zero) = No ESRD Y = Source of ESRD Variable is set to zero. If there is a start date before the diagnosis date at age 65 or older, or the last diagnosis if all are prior to turning age 65 then it is set to Y.
50	Chronic Renal Disease Year **(chr_esrd_yr)**	4	EDB	YYYY
54	Medicare Status Code **(med_stcd)**	2	EDB	Blank = Not Enrolled 10 = Aged 11 = Aged with ESRD 20 = Disabled 21 = Disabled with ESRD 31 = ESRD only Variable is selected by the first diagnosis date at age 65 or older, or the last diagnosis if all are prior to turning age 65. If information not available at those times then the first available data is used.
56	Medicare Status Code Year **(medst_yr)**	4	EDB	YYYY
60	Valid Date of Death **(vrfydth)**	1	EDB	N = No Y = Yes

Geographic Information
SEER program data are from the SEER record corresponding to the first diagnosis of cancer at age 65 or older. Zip codes are from Medicare enrollment file in the year of first diagnosis at age 65 or older or the last diagnosis if never 65. Census data was extracted from the 1990 (cen90) and 2000 (cen00) census data files for both census tract level information and zip code level information.

COL	FIELD	LENGTH	SOURCE	NOTES
61	State **(state)**	2	SEER	FIPS State Code
63	County **(county)**	3	SEER	FIPS County Code
66	Zip Code **(zip)**	9	EDB	Zip code. *** Special permission required for release.**

COL	FIELD	LENGTH	SOURCE	NOTES
75	Census Track Flag **(code_sys)**	1	SEER	Use in conjunction with Census Tract (tract) 0 = Not tracted 1 = 1970 Census tract definitions 2 = 1980 Census tract definitions 3 = 1990 Census tract definitions
76	Census Tract **(tract)**	6	SEER	1970/80/90 census tract *** Special permission required for release.**
82	Census Tract 2000 **(tract2k)**	6	SEER	2000 census tract *** Special permission required for release.**
88	HSA **(hsa)**	3	ARF	Health Service Area. Taken from the 2004 Area Resource File (ARF).
91	Urban/Rural recode **(urbrur)**	1	ARF	Urban/Rural Code 1 = Big Metro (Urban = 00 or 01) 2 = Metro (Urban = 02 or 03) 3 = Urban (Urban = 04 or 05) 4 = Less Urban (Urban = 06 or 07) 5 = Rural (Urban = 08 or 09) 9 = Unknown (Urban = 99)
92	Urban/Rural code **(urban)**	2	ARF	01-09, 99 see table at end of this document

Zip code and Census Tract level data

If there was no valid Zip Code reported in the Medicare file, zip code variables were left blank. If a zip code was reported in the Medicare file but not found in census data, the variables were 8-filled. Likewise if the area was not tracted (i.e., blank in the SEER record), census variables were left blank. If an invalid tract (i.e., 000000, 999999, 099999) was reported in SEER, the variables were 9-filled. If a tract was reported in SEER but not found in the census data the variables were 8-filled.

COL	FIELD	LENGTH	SOURCE	NOTES
94	Census Tract PCI **(ctpci90)**	10	Census 1990	Per Capita Income for census tract based on the 1990 Census Bureau survey.
104	Census Tract Median Income **(ctmed90)**	10	Census 1990	Median Income for census tract based on the 1990 census bureau survey.
114	Census Tract Density **(ctden90)**	10	Census 1990	Total pop/Total land area. Based on the 1990 census bureau survey.
124	Census Tract % Non HS **(ctnon90)**	6.2	Census 1990	Percent of persons 25+ with < 12 years education.

COL	FIELD	LENGTH	SOURCE	NOTES
130	Census Tract % HS only **(cthso90)**	6.2	Census 1990	Percent of persons 25+ with a high school education only.
136	Census Tract % Some College **(ctscl90)**	6.2	Census 1990	Percent of persons 25+ with some college education.
142	Census Tract % College **(ctcol90)**	6.2	Census 1990	Percent of persons 25+ with at least 4 years college education.
148	Census Tract % Black **(ctblk90)**	6.2	Census 1990	Black population / Total population
154	Census Tract % White **(ctwht90)**	6.2	Census 1990	White population / Total population
160	Census Tract % Hispanic **(cthsp90)**	6.2	Census 1990	Hispanic population / Total population
166	Census Tract % below poverty level **(ctpov90)**	6.2	Census 1990	Percent of tract residents living below the poverty level.
172	Census Tract % HHs speaking English not well or not at all **(cten590)**	6.2	Census 1990	Percent speaking English not well or not at all age 5+
178	Census Tract % HHs speaking English not well or not at all 65+ **(cten690)**	6.2	Census 1990	Percent speaking English not well or not at all age 65+

COL	FIELD	LENGTH	SOURCE	NOTES
184	Census Tract Median Income by Age **(medag90)**	10	Census 1990	Median HH Income by age. Census 1990 data not available.
194	Census Tract Median Income by Race **(medrc90)**	10	Census 1990	Median HH income by race. Census 1990 data not available.
204	Census Tract % Poverty by Race and Age **(povrc90)**	6.2	Census 1990	Percent of the tracts residents living below poverty level (Race and age specific <65, 65-74, 75+)
240	Census Tract % Non HS by Race **(nonrc90)**	6.2	Census 1990	Percent persons age 25+ with <12 years education. (Race specific)
216	Census Tract % HS only by Race **(hsorc90)**	6.2	Census 1990	Percent persons age 25+ with a high school diploma. (Race specific)
222	Census Tract $ Some College by Race **(sclrc90)**	6.2	Census 1990	Percent persons age 25+ with some college education. (Race specific)
228	Census Tract % College by Race **(colrc90)**	6.2	Census 1990	Percent persons age 25+ with at least 4 years of college education. (Race specific)
234	Census Tract % Poverty by Race **(pov_rc90)**	6.2	Census 1990	Percent of tract residents living below poverty level (Race specific)
240	Zip Code PCI **(zppci90)**	10	Census 1990	Per Capita Income for zip code.

COL	FIELD	LENGTH	SOURCE	NOTES
250	Zip Code Median Income **(zpmed90)**	10	Census 1990	Median HH Income for zip code.
260	Zip Code Density **(zpden90)**	10	Census 1990	Total pop/Total land area.
270	Zip Code % Non HS **(zpnon90)**	6.2	Census 1990	Percent of persons 25+ with < 12 years education.
276	Zip Code Tract % HS only **(zphso90)**	6.2	Census 1990	Percent of persons 25+ with a high school education only.
282	Zip Code % Some College **(zpscl90)**	6.2	Census 1990	Percent of persons 25+ with some college education.
288	Zip Code % College **(zpcol90)**	6.2	Census 1990	Percent of persons 25+ with at least 4 years college education.
294	Zip Code % Black **(zpblk90)**	6.2	Census 1990	Black population / Total population
300	Zip Code % White **(zpwht90)**	6.2	Census 1990	White population / Total population
306	Zip Code % Hispanic **(zphsp90)**	6.2	Census 1990	Hispanic population / Total population
312	Zip Code % HHs speaking English not well or not at all **(zpen590)**	6.2	Census 1990	Percent speaking English not well or not at all age 5+
318	Zip Code % HHs speaking English not well or not at all 65+ **(zpen690)**	6.2	Census 1990	Percent speaking English not well or not at all age 65+

COL	FIELD	LENGTH	SOURCE	NOTES
324	Zip Code % Non HS by Race **(znonrc90)**	6.2	Census 1990	Percent persons age 25+ with <12 years education. (Race specific)
330	Zip Code % HS only by Race **(zhsorc90)**	6.2	Census 1990	Percent persons age 25+ with a high school diploma. (Race specific)
336	Zip Code % Some College by Race **(zsclrc90)**	6.2	Census 1990	Percent persons age 25+ with some college education. (Race specific)
342	Zip Code % College by Race **(zcolrc90)**	6.2	Census 1990	Percent persons age 25+ with at least 4 years of college education. (Race specific)
348	Zip Code % Poverty by Race **(zppov_rc90)**	6.2	Census 1990	Zip Code Percent below poverty level by Race
354	Zip Code % Poverty by Race/Age **(zppovrage90)**	6.2	Census 1990	Zip Code Percent below poverty level by Race and Age
360	Zip Code % Poverty for All **(zppov90)**	6.2	Census 1990	Zip Code Percent below poverty level
366	Census Tract PCI **(ctpci00)**	10	Census 2000	Per Capita Income for census tract based on the 2000 Census Bureau survey.
376	Census Tract Median Income **(ctmed00)**	10	Census 2000	Median Income for census tract based on the 2000 census bureau survey.
386	Census Tract Density **(ctden00)**	10	Census 2000	Total pop/Total land area. Based on the 2000 census bureau survey.
396	Census Tract % Non HS **(ctnon00)**	6.2	Census 2000	Percent of persons 25+ with < 12 years education.

COL	FIELD	LENGTH	SOURCE	NOTES
402	Census Tract % HS only **(cthso00)**	6.2	Census 2000	Percent of persons 25+ with a high school education only.
408	Census Tract % Some College **(ctscl00)**	6.2	Census 2000	Percent of persons 25+ with some college education.
414	Census Tract % College **(ctcol00)**	6.2	Census 2000	Percent of persons 25+ with at least 4 Years college education.
420	Census Tract % Black **(ctblk00)**	6.2	Census 2000	Black population / Total population
426	Census Tract % White **(ctwht00)**	6.2	Census 2000	White population / Total population
432	Census Tract % Hispanic **(cthsp00)**	6.2	Census 2000	Hispanic population / Total population
438	Census Tract % below poverty level **(ctpov00)**	6.2	Census 2000	Percent of tract residents living below the poverty level.
444	Census Tract % HHs speaking English not well or not at all **(cten500)**	6.2	Census 2000	Percent speaking English not well or not at all age 5+
450	Census Tract % HHs speaking English not well or not at all 65+ **(cten600)**	6.2	Census 2000	Percent speaking English not well or not at all age 65+
456	Census Tract Median Income by Age **(medag00)**	10	Census 2000	Median HH Income by age. Census 1990 data not available.

COL	FIELD	LENGTH	SOURCE	NOTES
466	Census Tract Median Income by Race **(medrc00)**	10	Census 2000	Median HH income by race. Census 1990 data not available.
476	Census Tract % Poverty by Race and Age **(povrc00)**	6.2	Census 2000	Percent of the tracts residents living below poverty level (Race and age specific <65, 65-74, 75+)
482	Census Tract % Non HS by Race **(nonrc00)**	6.2	Census 2000	Percent persons age 25+ with <12 years education. (Race specific)
488	Census Tract % HS only by Race **(hsorc00)**	6.2	Census 2000	Percent persons age 25+ with a high school diploma. (Race specific)
494	Census Tract $ Some College by Race **(sclrc00)**	6.2	Census 2000	Percent persons age 25+ with some college education. (Race specific)
500	Census Tract % College by Race **(colrc00)**	6.2	Census 2000	Percent persons age 25+ with at least 4 years of college education. (Race specific)
506	Census Tract % Poverty by Race **(pov_rc00)**	6.2	Census 2000	Percent of tract residents living below poverty level (Race specific)
512	Zip Code PCI **(zppci00)**	10	Census 2000	Per Capita Income for zip code.
522	Zip Code Median Income **(zpmed00)**	10	Census 2000	Median HH Income for zip code.
532	Zip Code Density **(zpden00)**	10	Census 2000	Total pop/Total land area.

COL	FIELD	LENGTH	SOURCE	NOTES
542	Zip Code % Non HS **(zpnon00)**	6.2	Census 2000	Percent of persons 25+ with < 12 years education.
548	Zip Code Tract % HS only **(zphso00)**	6.2	Census 2000	Percent of persons 25+ with a high school education only.
554	Zip Code % Some College **(zpscl00)**	6.2	Census 2000	Percent of persons 25+ with some college education.
560	Zip Code % College **(zpcol00)**	6.2	Census 2000	Percent of persons 25+ with at least 4 years college education.
566	Zip Code % Black (**zpblk00**)	6.2	Census 2000	Black population / Total population
572	Zip Code % White **(zpwht00)**	6.2	Census 2000	White population / Total population
578	Zip Code % Hispanic **(zphsp00)**	6.2	Census 2000	Hispanic population / Total population
584	Zip Code % HHs speaking English not well or not at all **(zpen500)**	6.2	Census 2000	Percent speaking English not well or not at all age 5+
590	Zip Code % HHs speaking English not well or not at all 65+ **(zpen600)**	6.2	Census 2000	Percent speaking English not well or not at all age 65+
596	Zip Code % Non HS by Race **(znonrc00)**	6.2	Census 2000	Percent persons age 25+ with <12 years education. (Race specific)
602	Zip Code % HS only by Race **(zhsorc00)**	6.2	Census 2000	Percent persons age 25+ with a high school diploma. (Race specific)

COL	FIELD	LENGTH	SOURCE	NOTES
608	Zip Code % Some College by Race **(zsclrc00)**	6.2	Census 2000	Percent persons age 25+ with some college education. (Race specific)
614	Zip Code % College by Race **(zcolrc00)**	6.2	Census 2000	Percent persons age 25+ with at least 4 years of college education. (Race specific)
620	Zip Code % Poverty by Race **(zppov_rc00)**	6.2	Census 2000	Zip Code Percent below poverty level by Race
626	Zip Code % Poverty by Race/Age **(zppovrage00)**	6.2	Census 2000	Zip Code Percent below poverty level by Race and Age
632	Zip Code % Poverty for All **(zppov00)**	6.2	Census 2000	Zip Code Percent below poverty level
638	Filler	24		

Demographic Information
The following information comes from the SEER record corresponding to the first diagnosis at age 65 or over, or the latest diagnosis if all are prior to turning age 65.

COL	FIELD	LENGTH	SOURCE	NOTES
662	Sex **(s_sex)**	1	SEER	See Attachment A
663	Race Recode B **(rac_recb)**	2	SEER	Race Recode B from SEER 01,11 = White 01 = Caucasian, NOS 02 = Black 03 = American Indian/Alaska Native 04 = Chinese 05 = Japanese 06 = Filipino 07 = Hawaiian 08 = Other Asian or Pac. Islander 09 = Unknown 11 = Caucasian, Spanish origin or surname 12 = Other unspecified (1991+)
665	Race Recode Y **(rac_recy)**	1	SEER	See Attachment A

COL	FIELD	LENGTH	SOURCE	NOTES
666	ICD Code – Cause of Death **(icd_code)**	1	SEER	0 = Patient is alive at last follow-up 1 = Tenth ICD revision 8 = Eighth ICD Revision 9 = Ninth ICD revision
667	Cause of Death ICD -8 or 9 **(cod89v)**	4	SEER	0000 = Alive at last contact 7777 = State Death Certificate not available. 7797 = Death Certificate available, no COD listed.
671	Cause of Death ICD-10 **(cod10v)**	4	SEER	A020-Y891 0000 = Alive at last contact 7777 = State Death Certificate not available. 7797 = Death Certificate available, no COD listed.
675	Cause of Death to site recode KM **(codkm)**	3	SEER	This is a recode based on underlying cause of death to designate cause of death into groups similar to the incidence site recode with KS and mesothelioma. Study cutoff date has been applied, i.e. coded as alive if death occurred after study cutoff. Go to the end of this document for a listing of codes and their definition.
678	Cause of Death to site recode **(codpub)**	3	SEER	This recode was introduced to account for several newly valid ICD-10 codes and includes both cancer and non-cancer causes of death. Go to the end of this document for a listing of codes and their definition.
681	Filler	3		
684	NHIA Derived Hispanic Origin **(nhiade)**	1	SEER	See Attachment A
685	IHS Link Recode **(ihslink)**	1	SEER	See Attachment A
686	SEER Month of Death **(ser_dodm)**	2	SEER	00 = Alive 14 = Unknown month Date complete through 12/31/2010
688	SEER Year of Death **(ser_dody)**	4	SEER	0000 = Alive 2053 = Unknown year Date complete through 12/31/2010

COL	FIELD	LENGTH	SOURCE	NOTES
693	SEER Race/Ethnicity **(srace)**	2	SEER	See Attachment A
695	SEER Hispanic Surname **(origin)**	1	SEER	See Attachment A
696	Origin recode NHIA (Hispanic, Non-Hisp) **(origrecb)**	1	SEER	See Attachment A
697	Filler	10		
707	Vital Status Recode **(stat_rec)**	1	SEER	See Attachment A
708	Census Tract Certainty **(cen_cert)**	1	SEER	Associated with 1970/80/90 Census data 1 = Census tract based on complete and valid street address of residence 2 = Census tract based on residence ZIP + 4 3 = Census tract based on residence ZIP + 2 4 = Census tract based on residence ZIP code only 5 = Census tract based on ZIP code of post office box 6 = Census tract/BNA based on residence city where city has only one census tract, or based on residence ZIP code where ZIP code has only one census tract 9 = Unable to assign census tract based on available information

COL	FIELD	LENGTH	SOURCE	NOTES
709	Census Tract Certainty 2000 **(ctcer2k)**	1	SEER	Associated with 2000 Census data 1 = Census tract based on complete and valid street address of residence 2 = Census tract based on residence ZIP + 4 3 = Census tract based on residence ZIP + 2 4 = Census tract based on residence ZIP code only 5 = Census tract based on ZIP code of post office box 6 = Census tract/BNA based on residence city where city has only one census tract, or based on residence ZIP code where ZIP code has only one census tract 9 = Unable to assign census tract based on available information
710	SEER Year of Birth **(yr_brth)**	4	SEER	See Attachment A
714	Place of Birth **(plc_brth)**	3	SEER	See Attachment A
717	Number of SEER records **(count)**	2	Created at IMS	Number of eligible SEER records. SEER records prior to 1992 for Rural Georgia and records prior to 2000 for Greater California, New Jersey, Kentucky, and Louisiana were not included in PEDSF.
719	Number of primaries **(numprims)**	2	SEER	See Attachment A
722	Diagnosis indicator **(resnrec)**	1	Created at IMS	0 = Last Dx; Patient always less than 65. 1 = First Dx at age 65 or later.
729	Old Encrypted Casenum **(oshufcase)**	8	IMS	The encrypted case number for Louisiana and Seattle in the 2008 linkage.
737	Filler	12		

Repeated Part D Denominator Information

These variables (columns 750-829) are repeated once for every year from 2006 to 2011. (YY = the year of the file, ex. 06)

749	Research Triangle Institute Race Code (**rtirace**)	1	Part D Denom	Occurs just once. Taken from the last available record in the Part D Denominator files. Enhanced race/ethnicity designation based on first and last name algorithms. X = Enrolled in Medicare A and/or B, but no MIIR record found; unable to determine RTI Race Code 0 = Unknown 1 = Non-Hispanic White 2 = Black (or African American) 3 = Other 4 = Asian/Pacific Islander 5 = Hispanic 6 = American Indian/Alaska Native
750	On/Off Creditable Coverage Switch (**credYY**)	1	Part D Denom	Indicates for the Denominator reference year, the presence or absence of creditable coverage status. * = Enrolled in Medicare A and /or B, but no MIIR record for the year 0 = No instances of any creditable coverage status switch being ON at any point during the year. 1 = For at least 1 month during the year, 1 out of 5 creditable coverage switches was ON. Therefore, the beneficiary was enrolled in at Least 1 of 5 creditable coverage categories (i.e., FEHB, Tricare, VA, SPAP, or working aged).

751	Plan-Value Indicators (Jan. – Dec.) **(planYY_01-planYY_12)**	12	Part D Denom	12 monthly 1-byte indicators

Indicates for each month of the Denominator reference year, the Part D enrollment, which is based on the 1st digit or the Part D contract number.

Recodes only indicating type of plan; no 5-digit plan numbers.
0 = Not Medicare enrolled for the month
* = Enrolled in Medicare A and/or B, but no MIIR record for the month
H = Managed Care Organizations other than Regional PPO
R = Regional PPO
S = PDP
N = Not Part D Enrolled
E = Employer-sponsored (starting in Jan. 2007)

763	Denominator Cost Share Group (Jan. – Dec.) **(costYY_01-costYY_12)**	24	Part D Denom	12 monthly 2-byte indicators

Calculated field that describes a beneficiary's subsidy and/or copayment status.

00 = Not Medicare enrolled for the month
** = Enrolled in Medicare A and/or B, but no MIIR record for the month
01 = Bene is deemed with 100% premium-subsidy and no copayment
02 = Bene is deemed with 100% premium-subsidy and low copayment
03 = Bene is deemed with 100% premium-subsidy and high copayment
04 = Bene with LIS, 100% premium-subsidy and high copayment
05 = Bene with LIS, 100% premium-subsidy and 15% copayment
06 = Bene with LIS, 75% premium-subsidy and 15% copayment
07 = Bene with LIS, 50% premium-subsidy and 15% copayment
08 = Bene with LIS, 25% premium-subsidy and15% copayment
09 = No premium-subsidy and no copayment
10 = Not enrolled in Part D, but employer is entitled for RDS subsidy
11 = Bene with creditable coverage but no RDS

12 = Not Part D enrolled. No RDS and no
creditable coverage
13 = None of the above conditions have been

787	Retiree Drug Subsidy Indicators (Jan. – Dec.) **(rdsYY_01-rdsYY_12)**	12	Part D Denom	12 monthly 1-byte indicators

Indicates for each month of the Denominator reference year, whether the employer should be subsidized for the beneficiary.

0 = Not Medicare enrolled for the month
* = Enrolled in Medicare A and/or B, but no MIIR record for the month
Y = Employer subsidized for the retired beneficiary
N = No employer subsidization for the retired Beneficiary

799	State Reported Dual Eligible Status Code (Jan. – Dec.) **(dualYY_01-dualYY_12)**	24	Part D Denom	12 monthly 2-byte indicators

Indicates for each month of the Denominator reference year, the dual eligibility status, if any, for the beneficiary.

00 = Not Medicare enrolled for the month
** = Enrolled in Medicare A and/or B, but no MIIR record for the month
NA = Non-Medicaid
01 = QMB only
02 = QMB and Medicaid coverage including RX
03 = SLMB only
04 = SLMB and Medicaid coverage including RX
05 = QDWI
06 = Qualifying Individuals
08 = Other Dual Eligibles (Non-QMB, SLMB, QWDI, or QI) w/Medicaid coverage including RX
09 = Other Dual Eligibles but without Medicaid coverage
99 = Unknown

823	Plan Coverage Months **(ptdYY)**	2	Part D Denom	Contains the total number of months of Part D plan coverage for the beneficiary.
				The value in this field will be within the valid range of values 00 through 12, inclusive, dependent on the number of occurrences when the Plan Indicators = H, R, S, or E.
825	Retiree Drug Subsidy Months **(rdscntYY)**	2	Part D Denom	Contains the total number of months the employer is entitled to a retiree drug subsidy for the beneficiary.
				The value in this field will be within the valid range of values 00 through 12, inclusive, dependent on the number of occurrences when the Retiree Drug Subsidy Indicators = Y.
827	Medicaid Dual Eligible Months **(dualcntYY)**	2	Part D Denom	Contains the total numbers of months of dual eligibility for the beneficiary.
				The value in this field will be within the valid range of values 00 through 12, inclusive, dependent on the number of occurrences when the Medicaid Dual Eligible Indicators not equal to 00, **, NA, or 99.
829	Filler	1		

Repeated Entitlement Information

These variables (columns 1230-1273) are repeated once for every year from 1991 to 2011.

| 1230 | Entitlement Indicators **(mon1- mon252)** | 12 | EDB | One indicator for each month in the year.
'0' = Not entitled
'1' = Part A only
'2' = Part B only
'3' = Part A and B
Note: mon1 = 1/1991, mon2 = 2/1991, mon3 = 3/1991 ….mon252 = 12/2011 |

1242	HMO Indicators **(gho1-gho252)**	12	EDB	One indicator for each month in the year. 0 = Not a member of HMO 1 = Non-Lock-in, CMS to process provider claims 2 = Non-Lock-in, GHO to process in-plan Part A & in-area Part B claims 4 = Chronic care disease management organizations-FFS plan A = Lock-in, CMS to process provider claims B = Lock-in, GHO to process in-plan Part A & in-area Part B claims C = Lock-in, GHO to process all Part A and Part B claims Note: gho1 = 1/1991, gho2 = 2/1991, gho3 = 3/1991 gho252 = 12/2011
1254	State Buy-in Indicator **(allflag1-allflag252)**	12	EDB	One indicator for each month in the year. '0' = Not entitled '1' = Part A only '2' = Part B only '3' = Part A and B Note: allflag1 = 1/1991, allflag2 = 2/1991, allflag3 = 3/1991 ….allflag252 = 12/2011
1266	Part A Months **(ptacnt1991-ptacnt2011)**	2	EDB	Number of months covered for part A in each year covering 1991-2011. 00 = No coverage that year 01-12 = Number of months of coverage
1268	Part B Months **(ptbcnt1991-ptbcnt2011)**	2	EDB	Number of months covered for part B in each year covering 1991-2011. 00 = No coverage that year 01-12 = Number of months of coverage
1270	HMO Months **(hmocnt1991-hmocnt2011)**	2	EDB	Number of months as an HMO Member in each year covering 1991-2011. 00 = No coverage that year 01-12 = Number of months of coverage
1272	State Buy-in Months **(stbuyin1991-stbuyin2011)**	2	EDB	Number of months with state buy-in coverage In each year covering 1991-2011. 00 = No coverage that year 01-12 = Number of months of coverage

Repeated Zip Code and Registry Information

Zip Code and Registry (based on state and county) information has been retained from the EDB file for each year from 1991 to 2011.

2154	Zip Codes **(zip1991- zip2011)**	9	EDB	Zip code is assigned as the last zip code the patient resided in that year. ***Special permission required for release.**
2163	Registry Code **(registry1991- registry2011)**	2	EDB	01 = San Francisco

01 = San Francisco
02 = Connecticut
20 = Detroit
21 = Hawaii
22 = Iowa
23 = New Mexico
25 = Seattle
26 = Utah
27 = Atlanta
31 = San Jose
35 = Los Angeles
37 = Rural Georgia
41 = Greater California
42 = Kentucky
43 = Louisiana
44 = New Jersey
47 = Greater Georgia
NA = Has not lived in a registry yet
NOTE:
First registry found is retained until patient moved into a new registry.

2165	Registry Code #2 **(reg2cd1991-reg2cd2011)**	2	EDB	01 = San Francisco
				02 = Connecticut
				20 = Detroit
				21 = Hawaii
				22 = Iowa
				23 = New Mexico
				25 = Seattle
				26 = Utah
				27 = Atlanta
				31 = San Jose
				35 = Los Angeles
				37 = Rural Georgia
				41 = Greater California
				42 = Kentucky
				43 = Louisiana
				44 = New Jersey
				47 = Greater Georgia
				NA = Other

NOTE:
Registry is based off the state and county, which is the last residence the patient lived at for that year.

Repeated SEER Cancer Information

A patient can have up to 10 diagnoses in SEER, so the below variables are repeated for each cancer diagnosis the patient has in SEER. If the patient has more than 10 only the first 10 are retained for this file.
In parenthesis you will find the variable names. In italics you will find the SEER variable name.

2427	Sequence Number **(seq1-seq10)** *(seq_num)*	2	SEER	See Attachment A
2429	Primary Site **(site1-site10)** *(site02v)*	3	SEER	See Attachment A Note: This variable is length 4 in the SEER file. The first character (which is always 'C') was not written to the PEDSF.

2432	Site Recode ICD-O-3 **(siter1-siter10)** *(site_rec)*	2	SEER	A recode based on Primary Site and ICD-O-3 Histology in order to make analyses of site/histology groups easier. For example, the lymphomas are excluded from stomach and included with Lymphoma. Note: In this recode, Kaposi and mesothelioma are not separate categories based on histology. To match the site groupings in the CSR, do NOT use this recode -- use 'Site rec with Kaposi and mesothelioma'. Go to end of this document for a listing of codes and their definition.
2434	Site Recode with Kaposi and Mesothelioma ICD-O-3 **(siterkm1-siterkm10)** *(site2_re)*	2	SEER	A recode based on Primary Site and ICD-O-3 Histology in order to make analyses of site/histology groups easier. For example, the lymphomas are excluded from stomach and Kaposi and mesothelioma are separate categories based on histology. Go to end of this document for a listing of codes and their definition ***This variable is used to select patients for data requests.**
2438	Histology ICD-O-3 **(hist1-hist10)** *(hist03v)*	4	SEER	See Attachment A
2442	Histology ICD-O-2 (1992-2000) **(hist2_1-hist2_10)** *(hist02v)*	4	SEER	See Attachment A
2446	Behavior code ICD-O-3 **(beh1-beh10)** *(beh03v)*	1	SEER	See Attachment A
2447	Behavior code ICD-O-2 (1992-2000) **(beh2_1-beh2_10)** *(beh02v)*	1	SEER	See Attachment A

2448	Behavior recode for analysis **(behtrend1-behtrend10)** *(behanal)*	1	SEER	See Attachment A
2449	Month of Diagnosis **(modx1-modx10)** *(date_mo)*	2	SEER	The month of diagnosis is the month the tumor was first diagnosed by a recognized medical practitioner, whether clinically or microscopically confirmed. '01-'12' = Valid month '14' or '99' = Unknown
2451	Year of Diagnosis **(yrdx1-yrdx10)** *(date_yr)*	4	SEER	See Attachment A
2455	Diagnosis Date Flag **(ddxflag1-ddxflag10)**	2	SEER	12 = Date is unknown Missing = A valid date is provided, or the date was not expected to have been transmitted
2457	Age at Diagnosis recode **(ager1-ager10)** *(age_rec)*	2	SEER	See Attachment A
2459	Type of Reporting Source **(src1-src10)** *(rept_src)*	1	SEER	See Attachment A
2460	SEER Historic Stage A **(hstst1-hstst10)** *(hst_stga)*	1	SEER	See Attachment A
2461	SEER Modified AJCC Stage 3rd edition (1988-2003) **(aj3sr1-aj3sr10)** *(aj_3seer)*	2	SEER	See Attachment A

2463	AJCC Stage 3rd Edition (1988-2003) **(ajccstg1-ajccstg10)** *(ajcc_stg)*	2	SEER	See Attachment A
2465	Derived AJCC T, 6th edition (2004+) **(dajcct1-dajcct10)** *(d_ajcc_t)*	2	SEER	See Attachment A
2467	Derived AJCC M, 6th edition (2004+) **(dajccm1-dajccm10)** *(d_ajcc_m)*	2	SEER	See Attachment A
2469	Derived AJCC N, 6th edition (2004+) **(dajccn1-dajccn10)** *(d_ajcc_n)*	2	SEER	See Attachment A
2471	Derived AJCC Stage Group, 6th edition (2004+) **(dajccstg1-dajccstg10)** *(d_ajcc_s)*	2	SEER	See Attachment A
2473	Derived AJCC-Flag **(dajcflg1-dajcflg10)** *(d_ajcc_f)*	1	SEER	See Attachment A
2474	Derived SS1977 (2004+) **(dss77s1-dss77s10)** *(d_ssg77)*	1	SEER	See Attachment A

2475	Derived SS1977 - Flag **(dss77f1-dss77f10)** *(d_ssg77f)*	1	SEER	See Attachment A
2476	Derived SS2000 (2004+) **(dss00s1-dss00s10)** *(d_ssg00)*	1	SEER	See Attachment A
2477	Derived SS2000 – Flag **(dss00f1-dss00f10)** *(d_ssg00f)*	1	SEER	See Attachment A
2478	SEER Summary Stage 1977 (1995-2000) **(sss77v1-sss77v10)** *(ssg77)*	1	SEER	See Attachment A
2479	SEER summary stage 2000 (2001-2003) **(sssm2Z1-sssm2Z10)** *(ssg2000)*	1	SEER	See Attachment A
2480	Derived AJCC T Descriptor (2004+) **(t_dscrpt1-t_dscrpt10)** *(t_dscrpt)*	1	SEER	0 = Clinical stage 1 = Pathologic stage 2 = Autopsy stage 3 = Surgical resection performed, tumor size/extension based on 5 = Not Applicable
2481	Derived AJCC N Descriptor (2004+) **(n_dscrpt1-n_dscrpt10)** *(n_dscrpt)*	1	SEER	0 = Clinical stage 1 = Pathologic stage 2 = Autopsy stage 3 = Lymph nodes removed, evaluation based on pathologic evidence 5 = Not Applicable

2482	Derived AJCC M Descriptor (2004+) **(m_dscrpt1-m_dscrpt10)** *(m_dscrpt)*	1	SEER	0 = Clinical stage 1 = Pathologic stage 2 = Autopsy stage 3 = Pathologic examination, extension based on pathologic evidence 5 = Not Applicable
2483	Summary Stage 2000 (1998+) **(summ2k1-summ2k10)** *(hist_ssg_2000)*	1	SEER	See Attachment A
2486	First Malignant Primary Indicator **(frstprm1-frstprm10)** *(firstprm)*	1	SEER	See Attachment A
2488	Marital Status at Diagnosis **(marst1-marst10)** *(mar_stat)*	1	SEER	See Attachment A
2489	SEER Registry Code at Diagnosis **(reg1-reg10)** *(reg)*	2	SEER	01 = San Francisco (1973+) 02 = Connecticut (1973+) 20 = Detroit (1973+) 21 = Hawaii (1973+) 22 = Iowa (1973+) 23 = New Mexico (1973+) 25 = Seattle (1974+) 26 = Utah (1973+) 27 = Atlanta (1975+) 31 = San Jose (1988+) 35 = Los Angeles (1988+) 37 = Rural Georgia (1992+) 41 = Greater California (2000+) 42 = Kentucky (2000+) 43 = Louisiana (2000+) 44 = New Jersey (2000+) 47 = Greater Georgia (2000+)
2491	Month Therapy Started **(monrx1-monrx10)** *(monthrx)*	2	SEER	Month of Therapy 00 = No Therapy 01-12 = Valid month 14 = Unknown

2493	Year Therapy Started **(yearrx1-yearrx10)** *(year_rx)*	4	SEER	Year of Therapy. If year is blank, refer to Therapy Date Flag (dthflag) for more information.
2497	Radiation **(rad1-rad10)** *(radiatn)*	1	SEER	See Attachment A
2504	Other Therapy **(other_tx1-other_tx10)** *(othr_rx)*	1	SEER	0 = No other cancer-directed therapy 1 = Other cancer-directed therapy 2 = Other experimental cancer-directed therapy 3 = Double blind study, code not yet broken 6 = Unproven therapy (including Laetrile, Krebiozen, etc) 7 = Refused therapy 1-3 above 8 = Recommended, unknown if administered 9 = Unknown
2505	Radiation sequence with surgery **(radsurg1-radsurg10)** *(rad_surg)*	1	SEER	See Attachment A
2506	Radiation to Brain and /or CNS **(radbrn1-radbrn10)** *(rad_brn)*	1	SEER	See Attachment A
2507	Reason no cancer-directed surgery **(nosrg1-nosrg10)** *(no_surg)*	1	SEER	See Attachment A
2508	Grade **(grade1-grade10)** *(grade)*	1	SEER	See Attachment A

2509	Diagnostic Confirmation **(dxconf1-dxconf10)** *(dx_conf)*	1	SEER	See Attachment A
2510	Laterality **(lat1-lat10)** *(lateral)*	1	SEER	See Attachment A
2511	Site Specific Surgery (1983 – 1997) **(sssurg1-sssurg10)** *(ss_surg)*	2	SEER	See Attachment A
2513	Surgery of primary Site (1998-2002) **(sxprim1-sxprim10)** *(surgprim)*	2	SEER	See Attachment A
2515	Number of regional lymph nd exam (1998-2002) **(numnd1-numnd10)** *(surgnode)*	2	SEER	See Attachment A
2517	Scope of regional lymph node surgery (1998-2002) **(sxscop1-sxscop10)** *(scope02)*	1	SEER	See Attachment A
2518	Surgery of other reg/dist sites (1998-2002) **(sxsite1-sxsite10)** *(srgoth02)*	1	SEER	See Attachment A

2519	First course of reconstruction (1998-2002) **(recstr1-recstr10)** *(reconst)*	1	SEER	See Attachment A
2520	RX Summ – Surgery of Primary Site (1998+) **(sxprif1-sxprif10)** *(surgprim)*	2	SEER	See Attachment A
2522	RX –Summ Scope Reg LN Sur (2003+) **(sxscof1-sxscof10)** *(scope)*	1	SEER	See Attachment A
2523	RX Summ--Surg Oth Reg/Dis (2003+): **(sxsitf1-sxsitf10)** *(surgoth)*	1	SEER	See Attachment A
2524	Tumor Marker 1 (1990-2003) **(tumor1_1-tumor1_10)** *(tumor_1v)*	1	SEER	See Attachment A
2525	Tumor Marker 2 (1990-2003) **(tumor2_1-tumor2_10)** *(tumor_2v)*	1	SEER	See Attachment A
2526	Tumor Marker 3 (1998-2003) **(tumor3_1-tumor3_10)** *(tumor_3v)*	1	SEER	See Attachment A
2527	EOD 4 –size (1983-1987) **(e4siz1-e4siz10)** *(eod4)*	2	SEER	See Attachment A

2529	EOD 4 – extent (1983-1987) **(e4ext1-e4ext10)** *(eod4)*	1	SEER	See Attachment A
2530	EOD 4 – nodes (1983-1987) **(e4nod1-e4nod10)** *(eod4)*	1	SEER	See Attachment A
2531	EOD 10 - Tumor Size (1988-2003) **(e10sz1-e10sz10)** *(eod10_sz)*	3	SEER	See Attachment A
2534	EOD 10 – Tumor Extent (1988-2003) **(e10ex1-e10ex10)** *(eod10_ex)*	2	SEER	See Attachment A
2536	EOD 10 - Nodes (1988-2003) **(e10nd1-e10nd10)** *(eod10_nd)*	1	SEER	See Attachment A
2537	EOD 10 - # positive Nodes (1988+) **(e10pn1-e10pn10)** *(eod10_pn)*	2	SEER	See Attachment A
2539	EOD 10 - # Nodes Examined (1988+) **(e10ne1-e10ne10)** *(eod10_ne)*	2	SEER	See Attachment A
2541	EOD 10 – Prostate pathology ext (1995-2003) **(e10pe1-e10pe10)** *(eod10_pe)*	2	SEER	See Attachment A

2543	Coding system – extent of disease (1973-2003) **(eod_cd1-eod_cd10)** *(eodcode)*	1	SEER	See Attachment A
2544	County at Diagnosis **(cnty1-cnty10)** *(stcounty)*	3	SEER	FIPS County Code
2547	CS Extension (2004+) **(csex1-csex10)** *(cs_ext)*	3	SEER	See Attachment A
2550	CS Lymph Nodes (2004+) **(cslym1-cslym10)** *(cs_node)*	3	SEER	See Attachment A
2553	CS Mets at Dx (2004+) **(csmet1-csmet10)** *(cs_mets)*	2	SEER	See Attachment A
2555	CS Schema **(cssch1-cssch10)** *(csschema)*	2	SEER	See Attachment A
2557	CS Site-Specific Factor 1 (2004+) **(cs1st1-cs1st10)** *(cs_ssf1)*	3	SEER	See Attachment A
2560	CS Site-Specific Factor 2 (2004+) **(cs2st1-cs2st10)** *(cs_ssf2)*	3	SEER	See Attachment A
2563	CS Site-Specific Factor 3 (2004+) **(cs3st1-cs3st10)** *(cs_ssf3)*	3	SEER	See Attachment A

2566	CS Site-Specific Factor 4 (2004+) **(cs4st1-cs4st10)** *(cs_ssf4)*	3	SEER	See Attachment A
2569	CS Site-Specific Factor 5 (2004+) **(cs5st1-cs5st10)** *(cs_ssf5)*	3	SEER	See Attachment A
2572	CS Site-Specific Factor 6 (2004+) **(cs6st1-cs6st10)** *(cs_ssf6)*	3	SEER	See Attachment A
2575	CS Site-Specific Factor 8 (2004+) **(cs8st1-cs8st10)** *(cs_ssf8)*	3	SEER	See Attachment A
2578	CS Site-Specific Factor 10 (2004+) **(cs10st1-cs10st10)** *(cs_ssf10)*	3	SEER	See Attachment A
2581	CS Site-Specific Factor 11 (2004+) **(cs11st1-cs11st10)** *(cs_ssf11)*	3	SEER	See Attachment A
2584	CS Site-Specific Factor 13 (2004+) **(cs13st1-cs13st10)** *(cs_ssf13)*	3	SEER	See Attachment A
2587	CS Site-Specific Factor 15 (2004+) **(cs15st1-cs15st10)** *(cs_ssf15)*	3	SEER	See Attachment A
2590	CS Site-Specific Factor 16 (2004+) **(cs16st1-cs16st10)** *(cs_ssf16)*	3	SEER	See Attachment A

| 2593 | CS Site-Specific Factor 25 (2004+) **(cs25st1-cs25st10)** *(cs_ssf25)* | 3 | SEER | See Attachment A |
| 2596 | CS Tumor Size (2004+) **(cstum1-cstum10)** *(cs_size)* | 3 | SEER | See Attachment A |

| 2599 | CS Version 1st (2004+) **(csvf1-csvf10)** *(csv_org)* | 2 | SEER | This item indicates the number of the version used to initially code CS fields. The first two digits represent the major version number; the second two digits represent minor version changes; and, the last two digits represent even less significant changes, such as corrections of typographical errors that do not affect coding or derivation of results.
'00' = 000937
'01' = 010000
'02' = 010002
'03' = 010003
'04' = 010004
'05' = 010005
'06' = 010020
'07' = 010050
'08' = 010100
'09' = 010101
'10' = 010102
'11' = 010103
'12' = 010104
'13' = 010120
'14' = 010140
'15' = 010200
'16' = 010201
'17' = 010300
'18' = 010301
'19' = 010400
'20' = 010401
'21' = 014000
'22' = 014001
'23' = 014010
'24' = 020001
'25' = 020010
'26' = 020100
'27' = 020200
'28' = 020201
'29' = 020202
'30' = 020203
'31' = 020302 |

2601	CS Version Latest (2004+) **(csvl1-csvl10)** *(csv_der)*	2	SEER	This item indicates the number of the version of the CS used most recently to derive the CS output fields. The first two digits represent the major version number; the second two digits represent minor version changes; and, the last two digits represent even less significant changes, such as corrections of typographical errors that do not affect coding or derivation of results. '17' = 020100 '18' = 020200 '19' = 020302
2603	CS Tumor Size/Ext Eval (2004+) **(csts1-csts10)** *(exteval)*	1	SEER	See Attachment A
2604	CS Reg Nodes Eval (2004+) **(csrg1-csrg10)** *(nodeeval)*	1	SEER	See Attachment A
2605	CS Mets Eval (2004+) **(csmt1-csmt10)** *(metseval)*	1	SEER	See Attachment A
2606	Filler	2		
2608	Histology recode – Brain groupings **(hisrcb1-hisrcb10)** *(brainrec)*	2	SEER	See Attachment A
2610	ICD-O coding scheme **(icdo1-icdo10)** *(icdover)*	1	SEER	2 = Originally coded in ICD-O-2, (1973-2000) 3 = Originally coded in ICD-O-3, (2001+)

2611	NAPIIA Derived API Race **(napiia1-napiia10)** *(napiia)*	2	SEER	The NAACCR Asian Pacific Islander Identification Algorithm version 1 (NAPIIA v1) uses a combination of NAACCR variables to classify cases directly or indirectly as Asian Pacific Islander for analytic purposes. This version of the algorithm is focused on coding cases with a race code of Asian NOS (race code 96) to a more specific Asian race category, using the birthplace and name fields (first, last, and maiden names). Birthplace can be used to indirectly assign a specific race to one of eight Asian race groups (Chinese, Japanese, Vietnamese, Korean, Asian Indian, Filipino, Thai, and Cambodian). Names can beused to indirectly assign a specific race to one of seven Asian groups (Chinese, Japanese, Vietnamese, Korean, Asian Indian, Filipino, and Hmong). Future versions of NAPIIA will incorporate Pacific Islanders and will potentially incorporate name lists for Thai, Cambodian, and Laotians. Codes are the same as for Race/Ethnicity. Blanks are also in field and mean Algorithm not run .
2613	AYA site recode **(ayasiter1-ayasiter10)** *(aya_recode)*	2	SEER	See Attachment A
2615	Lymphoma subtype recode **(lymsubr1-lymsubr10)** *(lymphoma_recode)*	2	SEER	See Attachment A
2619	SEER cause-specific death classification **(vsrtdx1-vsrtdx10)** *(dth_class)*	1	SEER	See Attachment A

2624	Primary Payer at DX **(payer_dx1-payer_dx10)** *(payerdx)*	2	SEER	Primary Payer at Diagnosis identifies the patients primary insurance carrier or method of payment at the time of initial diagnosis and/or treatment.

01 = Not insured
02 = Not insured, self pay
10 = Insurance, NOS
20 = Private Insurance: Managed care, HMO, or PPO
21 = Private Insurance: Fee-for-Service
31 = Medicaid
35 = Medicaid - Administered through a Managed Care plan
60 = Medicare/Medicare, NOS
61 = Medicare with supplement, NOS
62 = Medicare - Administered through a Managed Care plan
63 = Medicare with private supplement
64 = Medicare with Medicaid eligibility
65 = TRICARE
66 = Military
67 = Veterans Affairs
68 = Indian/Public Health Service
99 = Insurance status unknown

2626	ER Status Recode Breast Cancer (1990+) **(erstat1-erstat10)** *(erstatus)*	1	SEER	See Attachment A
2627	PR Status Recode Breast Cancer (1990+) **(prstat1-prstat10)** *(prstatus)*	1	SEER	See Attachment A
2628	Lymph-vascular Invasion (2004+) **(vasinv1-vasinv10)** *(vasinv)*	1	SEER	See Attachment A

2629	Therapy date flag **(dthflag1- dthflag10)**	2	SEER	10 = Unknown if therapy was administered 11 = Therapy was not administered 12 = Therapy was administered, but the date is unknown Missing = A valid date is provided, or was not expected to have been transmitted
2631	Site-laterality-morph override **(ositmor1- ositmor10)** *(o_sitmor)*	1	SEER	See Attachment A
2632	Site-laterality-EOD override **(ositeod1- ositeod10)** *(o_siteod)*	1	SEER	See Attachment A
2633	Site-EOD-dx date override **(oeoddt1- oeoddt10)** *(o_eoddt)*	1	SEER	See Attachment A
2634	Site-behavior override **(ositbeh1- ositbeh10)** *(o_sitbeh)*	1	SEER	See Attachment A
2635	Surgery-diagnostic conf override **(osurcon1- osurcon10)** *(o_surcon)*	1	SEER	See Attachment A
2636	Site-type-lat-seq override **(oseqlat1- oseqlat10)** *(o_seqlat)*	1	SEER	See Attachment A
2637	Site-type edit override **(osittyp1- osittyp10)** *(o_sittyp)*	1	SEER	See Attachment A

2638	Sequence number-dx conf override **(oseqcon1-oseqcon10)** *(o_seqcon)*	1	SEER	See Attachment A
2639	Seq-ill-defined site override **(odfsite1-odfsite10)** *(o_dfsite)*	1	SEER	See Attachment A
2640	Report source sequence override **(orptsrc1-orptsrc10)** *(o_rptsrc)*	1	SEER	See Attachment A
2641	Leuk-Lymph dx confirmation override **(oleukdx1-oleukdx10)** *(o_leukdx)*	1	SEER	See Attachment A
2642	Histology edit override **(hbenign1-hbenign10)** *(h_benign)*	1	SEER	See Attachment A
2643	Age-site edit override **(ositage1-ositage10)** *(o_sitage)*	1	SEER	See Attachment A
2644	ICCC site recode ICD-O-3 **(ICCC3V1-ICCC3V10)** *(ICCC3)*	3	SEER	See Attachment A
2647	Histology recode - broad groupings **(histrec1-histrec10)** *(histrec)*	2	SEER	See Attachment A

2649	ICCC site recode extended ICD-O-3 **(ICCC3X1-ICCC3X10)** *(ICCC3EXT)*	3	SEER	See Attachment A
2652	ICCC site recode ICD-O-2 **(ICCCSITE1-ICCCSITE10)** *(ICCCSITE)*	3	SEER	See Attachment A
2655	SEER modified ICCC site recode ICD-O-2 **(ICCCSM1-ICCCSM10)** *(ICCCSM)*	3	SEER	See Attachment A
2658	Primary by International Rules **(intprim1-intprim10)** *(intprim)*	1	SEER	See Attachment A
2659	Survival time recode (month) **(survmo1-survmo10)** *(surv_tm)*	2	SEER	See Attachment A
2661	Survival time recode (year) **(survyr1-survyr10)** *(surv_tm)*	2	SEER	See Attachment A
2663	SEER other cause of death classification **(odthclass1-odthclass10)** *(o_dth_class)*	1	SEER	See Attachment A
2664	Type of follow-up expected **(typefu1-typefu10)** *(typefup)*	1	SEER	See Attachment A

2665	CS version input current (2004+) **(cscurrent1-cscurrent10)** *(csv_cur)*	2	SEER	This item indicates the number of the version of the CS after input fields have been updated or recoded. The first two digits represent the major version number; the second two digits represent minor version changes; and, the last two digits represent even less significant changes, such as corrections of typographical errors that do not affect coding or derivation of results. '17' = 020000 '21' = 020200 '22' = 020300 '23' = 020301 '24' = 020302
2667	CS Schema – v 0203 **(cs03sch1-cs03sch10)** *(cs0203schema)*	3	SEER	See Attachment A
2670	Recode ICD-O-2 to 9 **(icdot09_1-icdot09_10)** *(icdot09v)*	4	SEER	See Attachment A
2674	Recode ICD-O-2 to 10 **(icdot10_1-icdot10_10)** *(icdot10v)*	4	SEER	See Attachment A

SITE RECODE DEFINITIONS (SITER1-SITER10, SITE2r1-SITE2r10, SITERKM1-SITERKM10)

    1-10 = Oral cavity and pharynx
     01 = Lip
     02 = Tongue
     03 = Salivary gland
     04 = Floor of mouth
     05 = Gum & other mouth
     06 = Nasopharynx
     07 = Tonsil
     08 = Oropharynx
     09 = Hypopharynx
     10 = Other oral cavity and pharynx
    11-13,15-23,25-27,29-36 = Digestive system
     11 = Esophagus
     12 = Stomach
     13 = Small intestine
    15-23 = Colon exc. rectum
     15 = Cecum
     16 = Appendix
     17 = Ascending colon
     18 = Hepatic flexure
     19 = Transverse colon
     20 = Splenic flexure
     21 = Descending colon
     22 = Sigmoid colon
     23 = Large intestine, NOS
    25-26 = Rectum & rectosigmoid
     25 = Rectosigmoid junction
     26 = Rectum
    27 = Anus, anal canal & anorectum
    29-30 = Liver & Intrahepatic bile duct
     29 = Liver
     30 = Intrahepatic bile duct
     31 = Gallbladder
     32 = Other biliary
     33 = Pancreas
     34 = Retroperitoneum
     35 = Peritoneum, omentum & mesentery
     36 = Other digestive organs
    37-41 = Respiratory system
     37 = Nose, nasal cavity, middle ear
     38 = Larynx
     39 = Lung and bronchus
     40 = Pleura
     41 = Trachea, Mediastinum & Other respiratory organs

42 = Bones & joints
43 = Soft tissue (including heart)
44-45 = Skin (excluding basal and squamous)
  44 = Melanomas -- skin
  45 = Other non-epithelial skin
  46 = Breast
47-53 = Female genital system
  47 = Cervix
  48 = Corpus
  49 = Uterus, NOS
  50 = Ovary
  51 = Vagina
  52 = Vulva
  53 = Other female genital organs
54-57 = Male genital system
  54 = Prostate
  55 = Testis
  56 = Penis
  57 = Other male genital organs
58-61 = Urinary system
  58 = Bladder
  59 = Kidney and Renal pelvis
  60 = Ureter
  61 = Other urinary organs
  62 = Eye & orbit
63-64 = Brain and other nervous system
  63 = Brain
  64 = Cranial nerves Other nervous system
65-66 = Endocrine system
  65 = Thyroid
  66 = Other endocrine (including thymus)
68-69,71-72 = Lymphomas
68-69 = Hodgkins Disease
  68 = Nodal
  69 = Extranodal
71-72 = Non-Hodgkins lymphomas
  71 = Non-Hodgkins lymphomas - nodal
  72 = Non-Hodgkins lymphs-extranodal
  73 = Multiple myeloma
74-78,80,83,85,89 = Leukemias
74-76 = Lymphocytic
  74 = Acute lymphocytic
  75 = Chronic lymphocytic
  76 = Other lymphocytic
77-78,80,89 = Myeloid
  77 = Acute myeloid

    78 = Chronic myeloid
    80 = Acute monocytic leukemia
  83,85 = Other
    83 = Other acute leukemia
    85 = Aleukemic, subleukemic, and NOS
    86 = Miscellaneous
    89 = Other myeloid/monocytic leukemia
        87 = Mesothelioma (for siterkm1-siterkm10, only)
        88 = Kaposi Sarcoma (for siterkm1-siterkm10, only)

Codes for the variable CODPUB

    000 = Alive
    001 = Lip
    002 = Tongue
    003 = Salivary Gland
    004 = Floor of Mouth
    005 = Gum and Other Mouth
    006 = Nasopharynx
    007 = Tonsil
    008 = Oropharynx
    009 = Hypopharynx
    010 = Other Oral Cavity and Pharynx
    011 = Esophagus
    012 = Stomach
    013 = Small Intestine
    014 = Colon excluding Rectum
    024 = Rectum and Rectosigmoid Junction
    027 = Anus, Anal Canal and Anorectum
    029 = Liver
    030 = Intrahepatic Bile Duct
    031 = Gallbladder
    032 = Other Biliary
    033 = Pancreas
    034 = Retroperitoneum
    035 = Peritoneum, Omentum and Mesentery
    036 = Other Digestive Organs
    037 = Nose, Nasal Cavity and Middle Ear
    038 = Larynx
    039 = Lung and Bronchus
    040 = Pleura
    041 = Mediastinum and Other Respiratory Organs
    042 = Bones and Joints
    043 = Soft Tissue including Heart
    044 = Melanoma of the Skin
    045 = Other Non-Epithelial Skin

046 =Breast
047 =Cervix Uteri
048 =Corpus Uteri
049 =Uterus, NOS
050 =Ovary
051 =Vagina
052 =Vulva
053 =Other Female Genital Organs
054 =Prostate
055 =Testis
056 =Penis
057 =Other Male Genital Organs
058 =Urinary Bladder
059 =Kidney and Renal Pelvis
060 =Ureter
061 =Other Urinary Organs
062 =Eye and Orbit
090 =Brain and Other Nervous System
065 =Thyroid
066 =Other Endocrine including Thymus
067 =Hodgkin Lymphoma
070 =Non-Hodgkin Lymphoma
073 =Myeloma
074 =Acute Lymphocytic Leukemia
075 =Chronic Lymphocytic Leukemia
076 =Other Lymphocytic Leukemia
077 =Acute Myeloid Leukemia
080 =Acute Monocytic Leukemia
078 =Chronic Myeloid Leukemia
089 =Other Myeloid/Monocytic Leukemia
083 =Other Acute Leukemia
085 =Aleukemic, Subleukemic and NOS
086 =Miscellaneous Malignant Cancer
130 =In situ, benign or unknown behavior neoplasm
133 =Tuberculosis
136 =Syphilis
142 =Septicemia
145   Other Infectious and Parasitic Diseases including HIV
148 =Diabetes Mellitus
151 =Alzheimers (ICD-9 and 10 only)
154 =Diseases of Heart
157 =Hypertension without Heart Disease
160 =Cerebrovascular Diseases
163 =Atherosclerosis
166 =Aortic Aneurysm and Dissection
169 =Other Diseases of Arteries, Arterioles, Capillaries

172 =Pneumonia and Influenza
175 =Chronic Obstructive Pulmonary Disease and Allied Cond
178 =Stomach and Duodenal Ulcers
181 =Chronic Liver Disease and Cirrhosis
184 =Nephritis, Nephrotic Syndrome and Nephrosis
187 =Complications of Pregnancy, Childbirth, Puerperium
190 =Congenital Anomalies
193 =Certain Conditions Originating in Perinatal Period
196 =Symptoms, Signs and Ill-Defined Conditions
199 =Accidents and Adverse Effects
202 =Suicide and Self-Inflicted Injury
205 =Homicide and Legal Intervention
208 =Other Cause of Death
252 =State DC not available or state DC available but no COD

COD to site rec KM: (CODKM)

000 = Alive
001-014,024,027,029-062,065-067,070,
073-078,080,083,085-090 = All Malignant Cancers
001-010 = Oral Cavity and Pharynx
001 = Lip
002 = Tongue
003 = Salivary Gland
004 = Floor of Mouth
005 = Gum and Other Mouth
006 = Nasopharynx
007 = Tonsil
008 = Oropharynx
009 = Hypopharynx
010 = Other Oral Cavity and Pharynx
011-014,024,027,029-036 = Digestive System
011 = Esophagus
012 = Stomach
013 = Small Intestine
014,024 = Colon and Rectum
014 = Colon excluding Rectum
024 = Rectum and Rectosigmoid Junction
027 = Anus, Anal Canal and Anorectum
029-030 = Liver and Intrahepatic Bile Duct
029 = Liver
030 = Intrahepatic Bile Duct
031 = Gallbladder
032 = Other Biliary
033 = Pancreas
034 = Retroperitoneum

035 = Peritoneum, Omentum and Mesentery
036 = Other Digestive Organs
037-041 = Respiratory System
037 = Nose, Nasal Cavity and Middle Ear
038 = Larynx
039 = Lung and Bronchus
040 = Pleura
041 = Trachea, Mediastinum and Other Respiratory Organs
042 = Bones and Joints
043 = Soft Tissue including Heart
044-045 = Skin excluding Basal and Squamous
044 = Melanoma of the Skin
045 = Other Non-Epithelial Skin
046 = Breast
047-053 = Female Genital System
047 = Cervix Uteri
048-049 = Corpus and Uterus, NOS
048 = Corpus Uteri
049 = Uterus, NOS
050 = Ovary
051 = Vagina
052 = Vulva
053 = Other Female Genital Organs
054-057 = Male Genital System
054 = Prostate
055 = Testis
056 = Penis
057 = Other Male Genital Organs
058-061 = Urinary System
058 = Urinary Bladder
059 = Kidney and Renal Pelvis
060 = Ureter
061 = Other Urinary Organs
062 = Eye and Orbit
090 = Brain and Other Nervous System
065-066 = Endocrine System
065 = Thyroid
066 = Other Endocrine including Thymus
067,070 = Lymphoma
067 = Hodgkin Lymphoma
070 = Non-Hodgkin Lymphoma
073 = Myeloma
074-078,080,083,085,089 = Leukemia
074-076 = Lymphocytic Leukemia
074 = Acute Lymphocytic Leukemia
075 = Chronic Lymphocytic Leukemia

076 = Other Lymphocytic Leukemia
077-078,080,089 = Myeloid and Monocytic Leukemia
077 = Acute Myeloid Leukemia
080 = Acute Monocytic Leukemia
078 = Chronic Myeloid Leukemia
089 = Other Myeloid/Monocytic Leukemia
083,085 = Other Leukemia
083 = Other Acute Leukemia
085 = Aleukemic, Subleukemic and NOS
087 = Mesothelioma (ICD-10 only)
088 = Kaposi Sarcoma (ICD-10 only)
086 = Miscellaneous Malignant Cancer
130 = In situ, benign or unknown behavior neoplasm
133 = Tuberculosis
136 = Syphilis
142 = Septicemia
145 = Other Infectious and Parasitic Diseases including HIV
148 = Diabetes Mellitus
151 = Alzheimers (ICD-9 and 10 only)
154 = Diseases of Heart
157 = Hypertension without Heart Disease
160 = Cerebrovascular Diseases
163 = Atherosclerosis
166 = Aortic Aneurysm and Dissection
169 = Other Diseases of Arteries, Arterioles, Capillaries
172 = Pneumonia and Influenza
175 = Chronic Obstructive Pulmonary Disease and Allied Cond
178 = Stomach and Duodenal Ulcers
181 = Chronic Liver Disease and Cirrhosis
184 = Nephritis, Nephrotic Syndrome and Nephrosis
187 = Complications of Pregnancy, Childbirth, Puerperium
190 = Congenital Anomalies
193 = Certain Conditions Originating in Perinatal Period
196 = Symptoms, Signs and Ill-Defined Conditions
199 = Accidents and Adverse Effects
202 = Suicide and Self-Inflicted Injury
205 = Homicide and Legal Intervention
208 = Other Cause of Death
252 = State DC not available or state DC available but no COD

Note:

Rural/Urban Continuum as Defined in the 2004 ARF file

(urban/rural code)

The **2003 Rural/Urban Continuum Codes** are from Economic Research Service (ERS), Department of Agriculture. The codes form a classification scheme that distinguishes metropolitan (metro) counties by the population size of their metro area and nonmetropolitan (nonmetro) counties by degree of urbanization and adjacency to a metro area or nonmetro areas. All U.S. counties and county equivalents are grouped according to the official metro status announced by the Office of Management and Budget (OMB) in June 2003, when the population and worker commuting criteria used to identify metro counties were applied to results of the 2000 Census.

Metro counties are distinguished by population size of the Metropolitan Statistical Area of which they are part. Nonmetro counties are classified according to the aggregate size of their urban population. Within the three urban size categories, nonmetro counties are further identified by whether or not they have some functional adjacency to a metro area or areas. A nonmetro county is defined as adjacent if it physically adjoins one or more metro areas, and has at least 2 percent of its employed labor force commuting to central metro counties. Nonmetro counties that do not meet these criteria are classed as nonadjacent.

In concept, the 2003 version of the Rural-Urban Continuum Codes is comparable with that of earlier decades. However, OMB made major changes in its metro area delineation procedures for the 2000 Census, and the Census Bureau changed the way in which rural and urban are measured. Therefore, the new Rural-Urban Continuum Codes are not fully comparable with those of earlier years. OMBs changes added some additional metro areas by no longer requiring that a metro area must have at least 100,000 population if its urbanized area has no place of at least 50,000 people. More importantly, simplifying the worker commuting criteria that determine outlying metro counties had the effect of both adding numerous new outlying counties to metro status while deleting a smaller number that were previously metro.

The Census Bureau made a radical shift in determining rural-urban boundaries by changing and liberalizing the procedures for delineating urbanized areas of 50,000 or more people, and abandoning place boundaries in measuring urban or rural population. The procedures used in defining Urbanized Areas were extended down to clusters of 2,500 or more people, based solely on population density per square mile.

In earlier versions of the Rural-Urban Continuum Codes, metro areas with 1 million population or more were subdivided between central counties (Code 0) and fringe counties (Code 1). The Code 1 group has become much less meaningful in the last two censuses as more and more counties of large metro areas have been rated as central counties by OMB procedures. In 2000, only 1.6 percent of the population of large metro areas was in fringe counties. Therefore, this distinction has been dropped. Codes 0 and 1 have been combined, and the new code 1 represents all counties in metro areas of 1 million or more population.

The 2003 Rural/Urban Continuum Codes are defined as follows:

CODE   **CODEMETROPOLITAN COUNTIES (1-3)**
01   Counties of metro areas of 1 million population or more
02   Counties in metro areas of 250,000 - 1,000,000 population
03   Counties in metro areas of fewer than 250,000 population

NONMETROPOLITAN COUNTIES (4-9)
04 Urban population of 20,000 or more, adjacent to a metro area
05 Urban population of 20,000 or more, not adjacent to a metro area
06 Urban population of 2,500-19,999, adjacent to a metro area
07 Urban population of 2,500-19,999, not adjacent to a metro area
08 Completely rural or less than 2,500 urban population, adjacent to a metro area
09 Completely rural or less than 2,500 urban population, not adjacent to a metro area
99 Missing Value